Eur. J. Biochem. *239*, 220–228 (1996)

X-ray structure of yeast inorganic pyrophosphatase complexed with manganese and phosphate

Emil H. HARUTYUNYAN[1], Inna P. KURANOVA[1], Boris K. VAINSHTEIN[1], Wolfgang E. HÖHNE[2], Victor S. LAMZIN[3], Zbigniew DAUTER[3], Alexei V. TEPLYAKOV[3] and Keith S. WILSON[3]

[1] Institute of Crystallography, Russian Academy of Sciences, Moscow, Russia
[2] Institute of Biochemistry, Humboldt University, Berlin, Germany
[3] European Molecular Biology Laboratory (EMBL), c/o DESY, Hamburg, Germany

(Received 22 January 1996) – EJB 96 0082/3

The three-dimensional structure of the manganese-phosphate complex of inorganic pyrophosphatase from *Saccharomyces cerevisiae* has been refined to an *R* factor of 19.0% at 2.4-Å resolution. X-ray data were collected from a single crystal using an imaging plate scanner and synchrotron radiation. There is one dimeric molecule in the asymmetric unit. The upper estimate of the root-mean-square coordinate error is 0.4 Å using either the σ_A plot or the superposition of the two crystallographically independent subunits. The good agreement between the coordinates of the two subunits, which were not subjected to non-crystallographic symmetry restraints, provides independent validation of the structure analysis. The active site in each subunit contains four manganese ions and two phosphates. The manganese ions are coordinated by the side chains of aspartate and glutamate residues. The phosphate groups, which were identified on the basis of their local stereochemistry, interact either directly or via water molecules with manganese ions and lysine, arginine, and tyrosine side chains. The phosphates are bridged by two of the manganese ions. The outer phosphate is exposed to solvent. The inner phosphate is surrounded by all four manganese ions. The ion-binding sites are related to the order of binding previously established from kinetic studies. A hypothesis for the transition state of the catalytic reaction is put forward.

Keywords: pyrophosphatase; crystal structure; refinement; active site; manganese and phosphate ion binding.

Inorganic pyrophosphatases (PPases) belong to the family of phosphoryl-transfer enzymes. There are two groups of PPases that play different roles in cellular metabolism. Cytoplasmic PPases hydrolyse the high energy pyrophosphate (PP_i) bond thus removing this side product of biosynthetic reactions and controlling the level of pyrophosphate in the cell (Butler, 1971). Hydrolysis/synthesis of PP_i by membrane PPases is accompanied by proton transfer through the membrane and induction of membrane potential (Niren et al., 1984). PPases occupy a central role in intermediary metabolism.

More than 30 cytoplasmic and membrane PPases have been purified and characterised. The enzymes from the yeast *Saccharomyces cerevisiae* and from *Escherichia coli* have been most intensively studied (Cooperman, 1982; Avaeva and Nasarova, 1985; Lahti, 1983). They have very similar kinetic properties but different oligomeric structures. *S. cerevisiae* inorganic pyrophosphatase (YPPase) consists of two subunits of molecular mass 32 kDa, whereas the *E. coli* enzyme is a hexamer with subunits of 20 kDa (Hansen et al., 1972; Cohen et al., 1978; Josse and Wong, 1971). The sequence of YPPase has been determined (Cohen et al., 1978) and confirmed by gene sequencing (Kolakowski et al., 1988). Both *S. cerevisiae* and *E. coli* PPases are essentially perfect catalysts. The ratio k_{cat}/K_m (3×10^7 M^{-1} s^{-1}), which reflects enzyme efficiency, is close to the highest level possible for biochemical reactions under diffusion control (Fersht, 1984).

Bivalent metal cofactors are required for the activity of PPases. The efficiency of cations as activators decreases in the order: $Mg^{2+} > Zn^{2+} > Co^{2+} > Mn^{2+} > Cd^{2+}$. Ca^{2+} is a natural inhibitor (Butler, 1971). At equilibrium in the absence of metal, about 5% of the catalytic sites are occupied by PP_i (Janson et al., 1979). In the presence of metal ions, PPase catalysis is accompanied by oxygen exchange between phosphate and water.

Binding of metal ions to PPase and their role in catalysis have been intensively studied (Rapoport et al., 1973; Baykov and Avaeva, 1974; Cooperman et al., 1981; Knight et al., 1984). Biochemical, physicochemical, and kinetic studies of PPase were reviewed by Cooperman (1982) and a hypothetical scheme of PP_i hydrolysis proposed. The scheme is based on general base activation of an attacking nucleophilic water molecule and activation of the phosphoryl leaving group through metal ion complex formation and general acid catalysis.

A detailed kinetic scheme for YPPase catalysis was proposed by Baykov and Shestakov (1992), who suggested two possible catalytic pathways with participation of either three or four metal ions. The catalytic reaction involved several steps. Two metal ions bind first. Next, PP_i binds to the active site together

Correspondence to E. H. Harutyunyan, Institute of Crystallography, Russian Academy of Sciences, Leninsky pr. 59, Moscow 117333, Russia

Abbreviations. PPase, inorganic pyrophosphatase; YPPase, *Saccharomyces cerevisiae* inorganic pyrophosphatase; P_iL and P_iA, the two phosphate ions in the active site of inorganic pyrophosphatase; Mn1–Mn4, the four Mn^{2+} ions in the active site of inorganic pyrophosphatase.

Enzyme. Inorganic pyrophosphatase (EC 3.6.1.1).

Note. The coordinates of the model presented in this study together with the X-ray amplitudes and phases have been deposited in the Brookhaven Protein Data Bank (Bernstein et al., 1977) under accession codes 1YPP and R1YPPSF.

with the third (substrate) metal. The fourth metal ion binds either with very poor affinity to the complex of YPPase with two metals or more strongly to the complex of enzyme with two metals, PP_i and substrate metal. This provides further enhancement of the reaction. Kinetic constants were determined for all steps of the reaction with magnesium as cofactor.

Significant contributions to the understanding of YPPase arose from the determination of the crystal structures of the *S. cerevisiae* apoenzyme and its complexes with uranyl, terbium, or magnesium ions at 3-Å resolution and $CaPP_i$ at 5 Å (Kuranova et al., 1983; Terzyan et al., 1984). Even though the structure was not refined, the model revealed the essential architecture of the molecule and its secondary structure. The role of 17 amino acids in metal and substrate binding and a plausible scheme of catalysis were proposed (Kuranova, 1992). Chemical modification of YPPase confirmed the importance of Lys56, Tyr89, and Asp147 for catalysis (Komissarov et al., 1985; Gonzales and Cooperman, 1986; Raznikov et al., 1992). Recently, two more crystal structures have been solved for PPases from *Thermus thermophilus* (Teplyakov et al., 1994) and *E. coli* (Oganessyan et al., 1994; Kankare et al., 1994). Superposition of these on YPPase confirmed a high degree of similarity in the polypeptide fold, additional evidence of evolutionary divergence of the yeast and bacterial enzymes from a common ancestor.

The sequence alignment of PPases from different sources showed that they can be subdivided into two families (Cooperman et al., 1992). The identity between the most remote members of these families is less than 27%. All 17 residues implicated in the active site of YPPase from the X-ray structure are strictly conserved. There are only six conserved amino acids outside the active site. This supports a common catalytic mechanism for all these PPases. Site-directed mutagenesis of the *E. coli* enzyme confirmed several residues essential for activity (Lahti et al., 1990, 1991; Cooperman et al., 1992).

Determination of crystal structures of PPases complexed with functionally important ligands (P_i, PP_i analogues, and metals) is of particular value for a fuller understanding of the mechanism. The nature of the metal strongly influences substrate and product binding as well as the P_i/PP_i equilibrium. Thus phosphate binds much more strongly in the presence of manganese compared to magnesium (Cooperman et al., 1981). In this study, we have undertaken the structure analysis of the complex of YPPase with Mn^{2+} and P_i, the product of the reaction. Preliminary results have already been reported (Chirgadze et al., 1991). The refinement is now complete. Four Mn^{2+} and two P_i ions have been located in the active site.

MATERIALS AND METHODS

Enzyme isolation and purification. Enzyme purification was based on the procedure of Cooperman et al. (1973). The enzyme was assayed as described in Rapoport et al. (1972). The enzymatic activity was about 600–640 U/mg at room temperature.

Crystallisation. Crystals of the $Mn^{2+}P_i$/YPPase complex were grown from solutions of protein in 0.03 M Mes, pH 6.2, containing equal concentrations of both Mn^{2+} and phosphate (0.5 mM) and 2-methyl-2,4-pentanediol as precipitant as described in Chirgadze et al. (1991). The crystals were thick needles with a minimum dimension of 0.3 mm. They belong to the orthorhombic space group $P2_12_12_1$ with cell parameters a = 116.1, b = 106.2, and c = 56.1 Å. There is a dimer of the protein in the asymmetric unit which gives a V_m of 2.7 Da Å^{-3}.

X-ray data collection. X-ray intensities were collected at EMBL Hamburg using synchrotron radiation and an imaging plate scanner. The crystals did not diffract beyond 2.4 Å. 52 images with an oscillation range of 2° were recorded and processed using DENZO (Otwinowski, 1993). They were essentially complete at low resolution but only about 60% complete in the outer resolution shell (Table 1).

Table 1. Summary of data collection and processing. There is no σ cutoff in the number of unique reflections or the completeness.

Property	Value
Number of crystals	1
EMBL Hamburg beam line	X31
Wavelength (Å)	1.38
Resolution limit (Å)	2.4
Unique reflections	21 721
Redundancy	3.4
Overall completeness (%)	79
Overall $R(I)_{merge} = \Sigma\|I-\langle I\rangle\|/\Sigma I$ (%)	9.5
Completeness in the outer shell (%)[a]	62
$R(I)_{merge} = \Sigma\|I-\langle I\rangle\|/\Sigma I$ in the outer shell (%)[a]	19.9

[a] The outer shell corresponds to the 2.40–2.44-Å resolution range.

Table 2. Summary of refinement and characteristics of the model. All reflections were used with no σ cutoff.

Refinement characteristics		
Resolution range (Å)		10.0–2.4
Number of reflections		21 721
Number of protein atoms		4 496
Number of phosphate ions		4
Number of manganese ions		8
Number of water molecules		222
R factor (%)		19.0
σ_A estimate of root-mean-square coordinate error (Å)		0.37
Mean temperature factors (Å^2)		
Protein		29.5
Phosphate ions		26.7
Manganese ions		23.8
Water molecules		34.1
Root-mean-square deviation from ideality		
1–2 distances (Å)	0.020 (target)	0.022
1–3 distances (Å)	0.020 (target)	0.031
Temperature factor restraints		
Main chain 1–2 distances (Å^2)	2.0 (target)	2.6
Side chain 1–2 distances (Å^2)	4.0 (target)	6.4

Refinement. The initial model was that previously determined at 3.0-Å resolution for the apoenzyme (Kuranova et al., 1983; Terzyan et al., 1984) and preliminarily refined with the present data to an R factor of 25.4% (Chirgadze et al., 1991). In the present study, all X-ray data within the resolution range 10–2.4 Å were used from the start. First, the model was refined without stereochemical restraints using ARP (Lamzin and Wilson, 1993). This gave only marginal improvement in the density map, but allowed correction of some parts of the model. Further refinement was carried out using stereochemically restrained least-squares minimisation with the CCP4 (1994) version of PROLSQ (Konnert and Hendrickson, 1980) coupled with automated building and updating of the solvent structure again using ARP. Refinement was interspaced with manual correction using molecular graphics (Jones, 1978). The refinement is summarised in Table 2.

Identification of Mn^{2+}-binding and P_i-binding sites. The Mn^{2+} and P_i ions in the crystallisation liquor were expected to

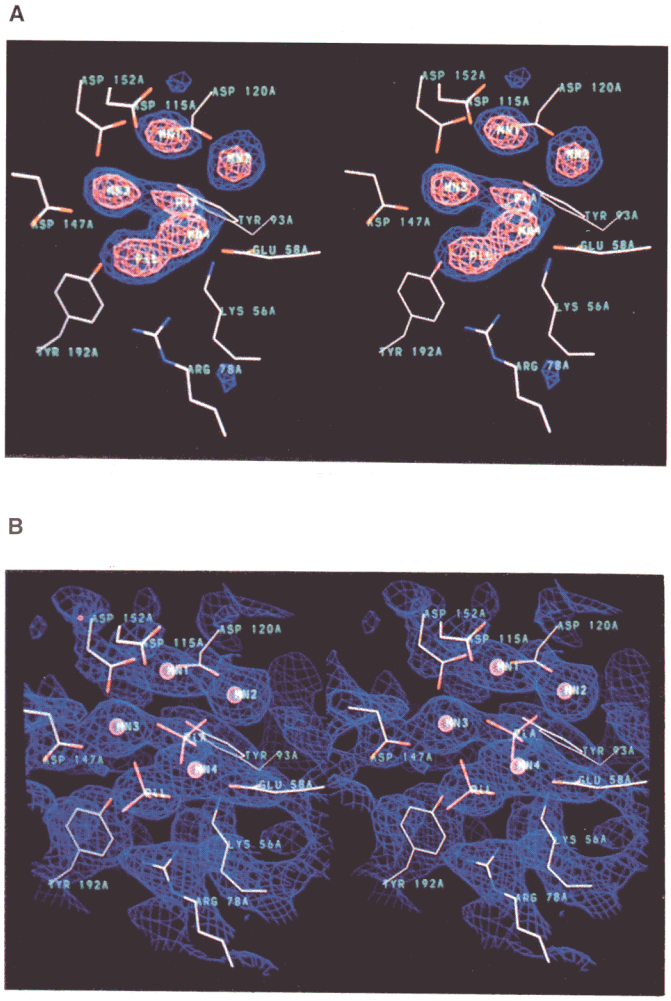

Fig. 1. Stereo view of the density map in the active site. Side chains interacting with Mn^{2+} and P_i groups (Table 3) are shown. (A) (F_o-F_c) map calculated with the Mn^{2+} and P_i not yet included in the model. Water molecules were accepted automatically as described in the text. Those in the active site, at that stage occupying the positions of the Mn^{2+} and P_i ions, were deleted and a further twenty cycles of least-squares minimisation were carried out to reduce the effects of memory before calculation of this map. The contours are blue (4 σ, 0.24 e Å^{-3}) and red (8 σ, 0.48 e Å^{-3}). (B) ($3F_o$-$2F_c$) map calculated with phases from the final model. The contour level is 1.5 σ (0.45 e Å^{-3}) above the mean.

bind to the active site. The electron density map, which was calculated when the refinement of the protein was deemed to have converged to an *R* factor of 21.1%, contained six large peaks in the active site in each of the two crystallographically independent subunits (Fig. 1A). Distances between the centres of these peaks were in the range 3.5−7.0 Å (Table 3). There were no significant deviations from the twofold non-crystallographic symmetry in spite of the fact that it was not imposed during refinement. Peaks in the two subunits were related with a root-mean-square deviation of 0.3 Å. The peak heights were similar and were two to three times higher than those for water molecules.

With data extending to atomic resolution (about 1.0 Å), the electron density at the centre of a phosphorus atom would be about 1.5 times lower than that at a Mn^{2+}. However, at 2.4-Å resolution both series termination ripples and overlap of the density for the oxygen and phosphorus atoms in P_i influence this ratio. As a result, electron density peaks corresponding to Mn^{2+} and P_i are of comparable heights (Table 3). Neither the height

Table 3. The Mn^{2+} and P_i peak heights and distances averaged for the two subunits. The ions were excluded from the phase calculation.

Peak no.	Site	Peak height ($3F_o$-$2F_c$) map	Distances (less than 4 Å) to other peaks (peak no. is given in parentheses)	Distances to nearest polar atoms of protein side chains
		e Å^{-3}	Å	
1	Mn1	2.9	3.8 (2) 3.8 (3) 3.5 (6)	2.4 (Asp115 OD1) 2.3 (Asp120 OD2) 2.0 (Asp152 OD1)
2	Mn2	2.8	3.8 (1) 3.4 (6)	2.0 (Asp120 OD1)
3	Mn3	2.5	3.8 (1) 3.6 (5) 3.3 (6)	2.2 (Asp147 OD1) 2.2 (Asp152 OD1)
4	Mn4	2.7	3.3 (5) 3.7 (6)	2.0 (Glu58 OE1)
5	P_iL	2.6	3.6 (3) 3.3 (4) 3.7 (6)	3.6 (Arg78 NH1) 3.6 (Arg78 NH2) 3.5 (Tyr192 OH)
6	P_iA	2.6	3.5 (1) 3.5 (2) 3.4 (3) 3.7 (4) 3.7 (5)	3.5 (Lys56 NZ) 3.6 (Tyr93 OH) 3.9 (Asp120 OD2)

nor the shape of the peaks allowed unambiguous identification. The atomic temperature factors can potentially be used to judge whether the identification of the peaks is correct. Unfortunately at limited resolution (2.4 Å), the refinement of individual atomic temperature factors (unless they are tightly restrained) is not justified. Thus, stereochemical insight was needed. There are only acidic residues around the first four peaks (Table 3). The distances between the centres of these peaks and carboxylate oxygens lie in the range 2.0−2.4 Å, which corresponds well to the typical distances around 2.2 Å observed in small molecules between Mn^{2+} and carboxylate ligands. These first four peaks can be unambiguously assigned to the four Mn^{2+} sites. The other two peaks are surrounded by arginine, lysine, and tyrosine side chains and the distances from the centres of the peaks to the nearest ligand atom are in the range 3.5−3.6 Å. They have been assigned as the two P_i groups.

Further refinement was carried out including the four Mn^{2+} and two phosphorus atoms in the model (Fig. 1B). Because of the limited resolution, the shape of the density for the P_i peaks was essentially spherical and the positions of P_i oxygen atoms could not be simply located from the density alone. The two phosphate sites are referred to as P_iL and P_iA. P_iL (peak 5) is located close to the Arg78 side chain. The phosphorus atom is 3.6 Å from the NH1 and NH2 atoms and lies in the plane of the guanidinium group. It is logical that two oxygen atoms of this P_i make two hydrogen bonds to the arginine NH atoms and lie in the same plane. The positions of all atoms in P_iL are defined based on the tetrahedral coordination of the phosphorus. Coordination of one P_i group by Arg78 was proposed earlier (Cooperman, 1982).

The CA atoms of nine residues involved in binding Mn^{2+} and P_i (Table 3) superimpose on the equivalent atoms in *T. thermophilus* PPase with a root-mean-square deviation of 1.5 Å. The location of P_iL is similar to that of the sulphate ion in the

active site of the *T. thermophilus* enzyme, which was attributed to the P_i-binding site. The distance between the P_iL phosphorus and the sulphur is 1.3 Å. This supports that peak 5 corresponds to P_i.

The peaks corresponding to P_iA (peak 6) in both subunits have two small protuberances, which we have assumed to represent the positions of two oxygens. The positions of the other two oxygens in P_iA were derived from these. However, the orientations of neither P_iA nor P_iL are defined with certainty in the present model.

RESULTS AND DISCUSSION

Quality of the model and non-crystallographic symmetry. Characteristics of the final model are given in Table 2. Only residues 1−282 in both subunits are visible in the electron density map. The remaining four C-terminal residues are disordered. Each subunit contains 2248 protein atoms, four Mn^{2+} ions, and two phosphate groups. There are 222 solvent molecules in the model, approximately 0.4/amino acid residue. This relatively small number is a consequence of the limited resolution of the X-ray data. The difference map calculated with final model phases has minimum, maximum, and root-mean-square values of −0.29, 0.51, and 0.05 e $Å^{-3}$, respectively, and does not contain significant features to suggest further improvement of the model. The Ramachandran plot calculated using PROCHECK (Morris et al., 1992) showed no residues with main chain dihedral angles in disallowed regions. 85% of the residues are in the core region.

The overall estimate of coordinate error from the σ_A plot (Read, 1986) is 0.37 Å. The two independent subunits in the crystallographic asymmetric unit and refinement of the model without non-crystallographic symmetry restraints allows an alternative means of estimating coordinate error. In the absence of crystal lattice contacts giving systematic deviations between non-crystallographically related subunits, the root-mean-square deviation between them, divided by $\sqrt{2}$, would approximately correspond to the coordinate error. Superposition of the two subunits results in a root-mean-square deviation of 0.36 Å for equivalent CA atoms, 0.38 for main chain and 0.88 for side chain atoms. The relatively high value obtained for the side chains is affected by disorder modelled differently in the two subunits. When 34 pairs of side chains with average temperature factors of more than 50 $Å^2$ were excluded, the root-mean-square deviation for the remaining 248 decreased to 0.59 Å. The disordered side chains are mainly lysines, arginines, and carboxylates on the protein surface. The active site Mn^{2+} atoms from the two subunits superimpose with a root-mean-square deviation of 0.16 and the P_i phosphorus atoms 0.23 Å. 57 pairs of solvent atoms (51% of the total solvent sites) closely follow non-crystallographic symmetry with a deviation of less than 1.0 Å.

The mean temperature factor and root-mean-square deviation from non-crystallographic twofold symmetry for the main chain atoms is plotted as a function of residue number in Fig. 2. The root-mean-square deviation is highly correlated with the temperature factor. A few segments of the backbone, the loops connecting strands or helices (Ala36 to Ile42, connecting β3′ and β4; Ala73 to Lys76, β6 and β7; Asn104 to Lys111, β10 and β11; Lys233 to Asp241, α3 and α4) deviate significantly from non-crystallographic symmetry, which is clearly a result of crystal lattice contacts. Thus, for atoms which are well ordered and not involved in lattice contacts, the estimate of coordinate error, obtained as the root-mean-square deviation from twofold symmetry divided by $\sqrt{2}$, is about 0.2 Å for main chain and 0.4 Å for side chain atoms. This is in a good agreement with the overall error estimate of 0.37 Å derived from the σ_A plot.

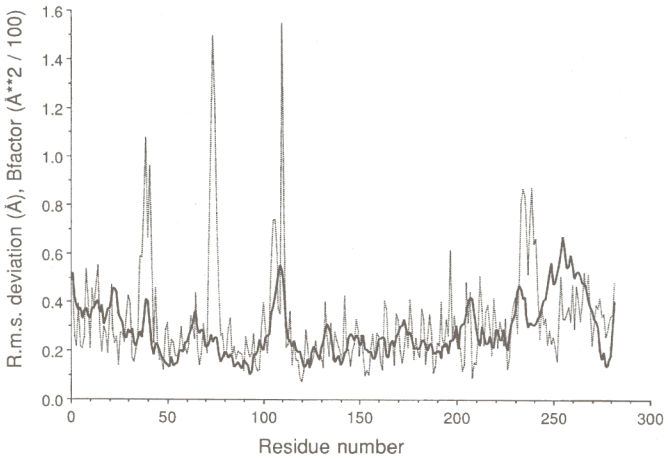

Fig. 2. Root-mean-square deviation from non-crystallographic twofold symmetry (——) and mean temperature factors (——) for main chain atoms as a function of residue number. R.m.s., root-mean-square.

The non-crystallographic twofold symmetry was not imposed during refinement. This allowed independent validation of the significance of the model as described above. No significant deviations from twofold symmetry were observed in the active site in terms of the P_i and Mn^{2+} positions. The only discrepancies lay in the presence or absence of weakly bound water molecules in the metal coordination spheres (see below). These do not affect the conclusions drawn in this study. Subsequently, 20 cycles of refinement were run with the imposition of non-crystallographic symmetry, with tight restraints on the main chain and weaker ones on the side chain atoms. The *R* factor increased slightly and there was little change in the density maps (data not shown). No new conclusions could be derived concerning the nature of the active site. The model deposited in the data bank and described elsewhere in the text is that without non-crystallographic symmetry restraints imposed.

The fold of YPPase. The YPPase molecule is made up of two subunits (Fig. 3). The two active sites are about 40 Å apart. Each subunit has a compact globular shape (Fig. 4), and the overall architecture of the initial apoYPPase model is retained, with differences in the definition of secondary structure elements. Some new short β-strands and α-helices have been identified in the present complex and some of those previously assigned in the apoenzyme have been removed. The topology is shown in Fig. 5, using the numbering given for the apoenzyme. Compared to the original structure, β-strands 8, 12, and 18 are no longer present; β-strands 15 and 16 comprise a single strand, β15; strands β3 and β7 are each split into two shorter strands and two new β-strands 18′ and 19 as well as the α0-helix have been identified.

The core of the subunit is formed by a five-stranded β-barrel (Figs 4 and 5) consisting of β-strands 11, 9, 4, 13, and the C-terminal part of β15. Four α-helices (α0, α1, α2, and α3) surround the barrel and are located at the protein surface. The arrangement of the β-barrel and two α-helices (α2 and α3) is conserved in other PPases with known three-dimensional structure. The second part of the subunit comprises an extensive β-sheet made up of four long (β1, β2, β6, and β5) and three short (β18′, β7, and β3) β-strands. The β-sheet makes contacts with two short β-strands at the C-terminus (β19 and β7′) and is shielded from the solvent by the α4-helix (Figs 4 and 5) as is the β-barrel. Although the YPPase subunit clearly consists of the two parts, β-barrel and β-sheet, they pack closely together and do not form

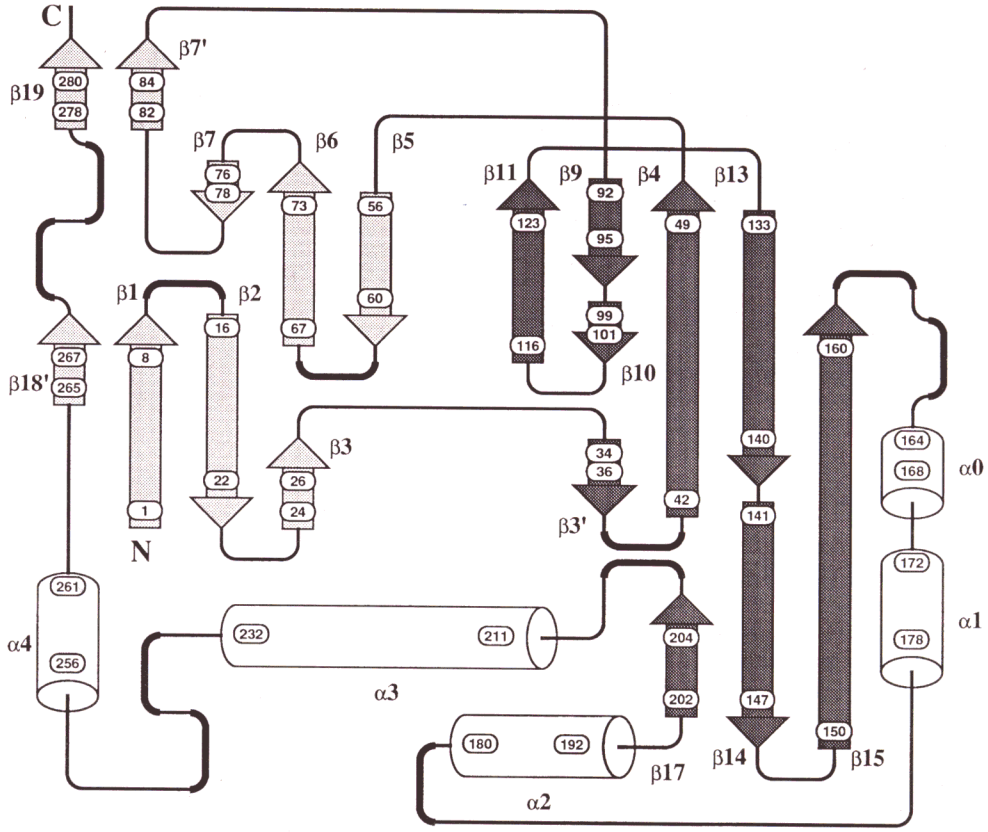

Fig. 5. Schematic representation of the YPPase topology. β-Strands are shown as arrows, α-helices as cylinders, β-turns as thick curved line. The residue numbers indicate the boundaries of the secondary structure elements. The β-strands of the β-barrel are shown in dark grey.

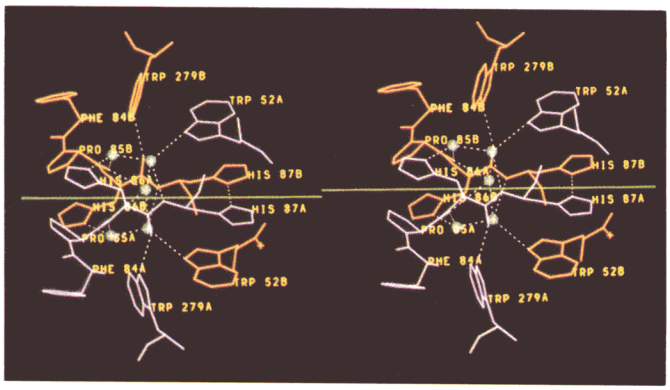

Fig. 6. Stereo view of the central part of the intersubunit interface. Water molecules are shown by dotted spheres. The two-fold molecular symmetry axis is shown in green.

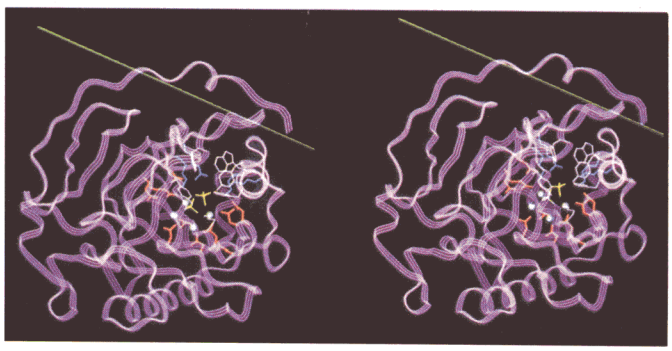

Fig. 7. Stereo view of the active site looking down into the cavity. The YPPase subunit is shown in magenta. The twofold molecular symmetry axis and P_i groups are in green and the Mn^{2+} ions are dotted spheres. Three groups of residues composing the active site are the hydrophobic core comprising mostly aromatic residues at the bottom of the cleft in white, carboxylates from the large β-sheet and β-barrel in red, and positively charged residues covering the hydrophobic core in cyan.

atom of each of three carboxylate side chains from Asp115, Asp120, and Asp152. It also coordinates a P_iA oxygen. In subunit A, two water molecules complete the octahedral coordination of Mn1. In subunit B, these water molecules are not observed. Mn2 is in contact with the second oxygen atom of the bridging Asp120 and a P_iA oxygen. Four water molecules complete the octahedral coordination of Mn2 in the A subunit but only one water molecule is observed in the B subunit. The Mn1 and Mn2 ions together with lysine and arginine side chains create a positively charged environment favouring binding of the two P_i groups complexed with the other two Mn^{2+} ions.

The second pair of metal ions, Mn3 and Mn4, act as bridges between the two phosphates. Mn3 coordinates oxygen atoms of the Asp147 and Asp152 carboxyl groups and oxygens of P_iA and P_iL. Two water molecules are coordinated with Mn3 in subunit B but only one in subunit A. Mn4 has square pyramidal coordination in both subunits. Besides the two oxygen atoms of the P_iA and P_iL groups, it is bound to an oxygen atom of the Glu58 carboxylate group and two water molecules. The square base of the pyramid is exposed to solvent. The coordination of all Mn^{2+} ions is probably in reality octahedral, with the apparently vacant positions occupied by mobile water molecules with too high temperature factors to be observed in the present analysis. The mean values of distances from Mn^{2+} ions to their ligands are as follows: O (aspartate or glutamate) 2.1 Å, O (P_i) 2.3 Å, and water molecules 2.2 Å.

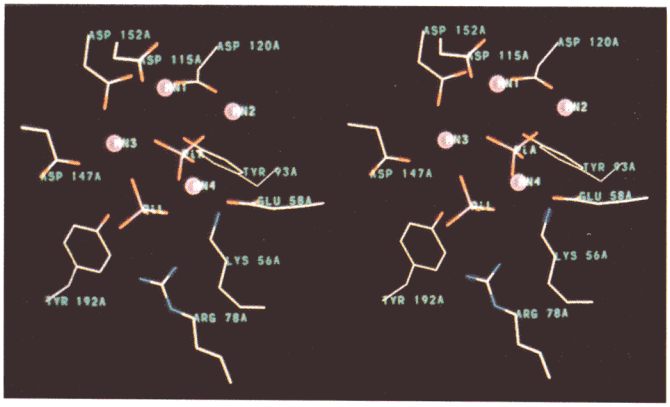

Fig. 8. Stereo view of the active site residues directly involved in binding Mn^{2+} (red dotted spheres) and P_i ions.

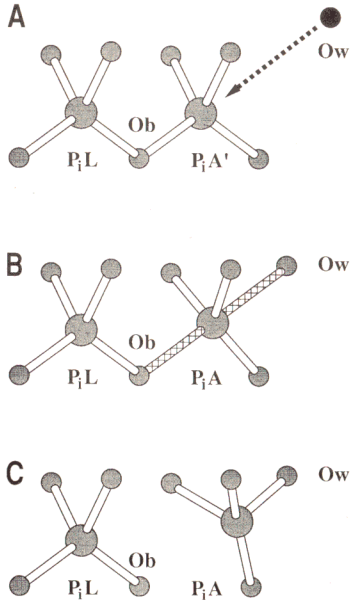

Fig. 9. Scheme for formation of the transition state.

P_i groups interact with side chains that belong to either the positively charged or hydrophobic clusters in the active site. The distance between the phosphorus atoms of the two P_i groups is 3.7 Å and 3.8 Å in the two subunits. Thus, there appears to be no pyrophosphate but rather two separate P_i groups in the complex. P_iA is located deep in the cavity and is partially shielded from solvent by P_iL. P_iA interacts with all four metal ions and with Lys56, Tyr93, Tyr192, and Lys193. P_iL bridges two Mn^{2+} ions and is hydrogen bonded to Arg78 NH1 and NH2, Lys193 NZ and Tyr192 OH.

Mn^{2+} and P_i binding. The reaction catalysed by PPases is dependent on the presence of a natural metal activator, Mg^{2+}. Metal activators bind to the active site successively and the dissociation constant for the first Mg^{2+} (0.01 mM) is much lower than that for the second (0.22 mM). The third Mg^{2+} binds much more weakly (27 mM; Baykov and Shestakov, 1992). Binding of PP_i or P_i complexed with the third metal ion facilitates binding of a fourth metal ion at the same concentration as that for the second. It is likely that the presence of only three metal ions in the active site is sufficient for catalysis to proceed. However, the catalytic reaction speeds up in the presence of the fourth ion. The reaction

$$MgP_2O_7^{2-} + Mg^{2+} + H_2O \rightleftharpoons 2MgPO_4^- + 2\,H^+$$

involves proton release without protonation of the leaving phosphate groups.

Of the four manganese ions, Mn1 is buried most deeply in the active site and is shielded from solvent by P_iA. Mn1 can be assumed to be the highest affinity metal ion, i.e. the first metal to bind to the active site. Stereochemical evidence, i.e. that Mn1 is bound by three aspartate groups, supports this assumption. In the structure of *E. coli* PPase with a low concentration of manganese (Harutyunyan et al., 1996), only one Mn^{2+} is present and occupies the position corresponding to Mn1.

Mn4 is the metal site most accessible to solvent and is likely to be the fourth manganese that binds most weakly and is not absolutely required for catalysis. Binding of Mn4 corresponds to the second alternative pathway for the reaction, described by Baykov and Shestakov (1992).

The apparent accessibility of the metal-ion-binding sites alone does not clearly identify the metal that binds second: Mn2 or Mn3. From stereochemical considerations only (binding by two aspartate residues), Mn3 could be considered as the second most tightly bound metal. However, Mn3 but not Mn2 is in a position to bridge the two phosphate moieties in the PP_i group and hence must be the so-called substrate metal that is third in order of binding. Therefore, Mn2 corresponds to the second metal.

Of the two phosphate ions, P_iA is most deeply buried and is expected to be the first phosphate to bind during the reverse reaction, when the substrate metal Mn3 and P_iA are believed to bind essentially simultaneously to the dimetal complex. P_iA is liganded by all four Mn^{2+} ions and its P-O bonds should be weakened more than those in P_iL which only interacts with two metals. Therefore, the P_iA phosphorus is likely to be that on which displacement occurs during catalysis in both the forward and inverse reactions.

P_iL corresponds to the 'entering' P_i, the second phosphate to bind in the inverse reaction and the first leaving group in the forward reaction. In summary, the last groups to bind in the inverse reaction and the first to leave in the forward reaction are likely to be P_iL and Mn4, which are most exposed to solvent.

The presence of two P_i groups and four Mn^{2+} ions indicates that the crystal contains the product complex of PP_i hydrolysis. This is confirmed by an additional observation on a crystalline complex of YPPase with $Mn(PO_3)_2NH$ (Kuranova, I. P. and Teplyakov, A. V., unpublished results) where the active site contained six peaks in the same positions as in the present structure, i.e. the products of $(PO_3)_2NH$ hydrolysis occupy the same positions as the present P_i groups. The structural data on the present Mn^{2+}-P_i YPPase complex clarify some aspects of the catalytic mechanism and of the transition state.

Mechanism. Hydrolysis of PP_i proceeds according to an S_N2 mechanism (Gonzales et al., 1986). A hydroxyl anion, Ow, attacks the P_iA phosphate group in the direction opposite to the Ob ether P-O bond (Fig. 9A). Trigonal bipyramidal coordination of the P_iA phosphorus in the transition state is stabilised by metal ions and positively charged protein side chains, the combined effects of which weaken the P-O bonds. Due to these interactions, the phosphorus atom of P_iA is displaced towards Ow (Fig. 9B). Such displacement (about 1.0 Å in total) leads to rupture of the P-Ob bond (Fig. 9C). The inverse reaction, PP_i synthesis from two P_i groups (P_iL and P_iA'), involves protonation of the Ow oxygen on the P_iA' group. The P_iL group via the Ob oxygen attacks the phosphorus of the P_iA. Displacement of the phosphorus towards P_iL Ob and release of a water molecule

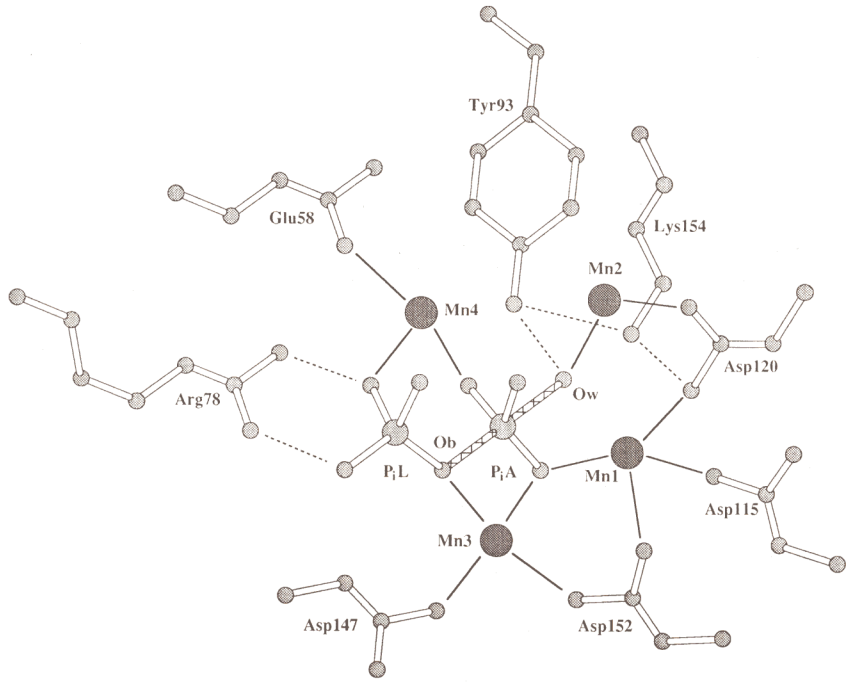

Fig. 10. Schematic representation of the proposed transition state. Hydrogen bonds between residues presumed to be involved in the activation of the water molecule are shown.

leads to formation of PP_i. The same protein side chains and metal ions are presumably involved in both the forward and reverse reactions. The proposed mechanism implies firstly that Ow (P_iA') should be most remote from P_iL and secondly that the Ob of P_iL and the phosphorus and Ow of P_iA' should lie on a straight line in the transition state and approximately in a straight line during the subsequent stage of hydrolysis when PP_i is transformed into two separate tetrahedral P_i groups.

The hypothetical transition state is shown in Fig. 10. The leaving P_iL is coordinated by Mn3 and Mn4 and anchored by Arg78. The scheme involves the movement of P_iA from its position in the present crystal structure to a new position P_iA' where it is covalently linked with Ob. The P_iA phosphorus shifts from the starting position towards Ob to become five-coordinate. P_iA interacts with all four Mn^{2+} ions. This is the P_i that enters and leaves the active site complexed with the substrate metal Mn3. In the transition state, Mn3, which bridges two isolated P_i groups in the crystal structure, is assumed not to change its position and to interact with two P_iA oxygen atoms one of which is Ob. This introduces tension in the transition state structure since the distance between the positively charged Mn3 and the phosphorus of P_iA is short. However, such tension promotes movement of the phosphorus atom towards Ow during hydrolysis of PP_i.

As the positions of the P_iA oxygen atoms are not determined reliably in the present study, it is not possible to define which is Ow and, as a consequence, which is Ob in P_iL. However, in P_iA the Ow oxygen atom should be that most remote from P_iL, i.e. in the region close to Mn2, which is probably involved in the activation of the water molecule. Asp120, which bridges Mn1 and Mn2, forms a hydrogen bond to Lys154 NZ, which has a hydrogen bond to Tyr93 OH. These residues together with Mn^{2+} ions are probably involved in deprotonation of the catalytic Ow and protonation of the corresponding oxygen atom of P_iA.

The structure of the transition state deduced from the present crystal structure agrees with the overall scheme suggested by Cooperman et al. (1992). However, there are significant differences. Binding of PP_i by two bridging metals was suggested by Kuranova and Sokolov (1986). According to the present scheme, the second bridging metal is not the second metal-activator but rather Mn4 which enters the active site last.

CONCLUSIONS

In discussing the catalytic mechanism, it has been assumed that only minor conformational changes occur when the transition state is formed. Further structures of crystalline PPase complexed with metal ions with and without substrate analogues are essential and will allow a more detailed understanding of the subsequent stages of the reaction. Some crystalline complexes of YPPase have already been obtained but their diffraction extended to less than 3-Å resolution (Kuranova et al., 1990). Such studies are now in progress. Different complexes of both yeast and *E. coli* PPases diffracting to about 2.0-Å resolution have already been crystallised (Heikineimo et al., 1995). Crystalline complexes of YPPase complexed both with two Mn^{2+} ions and ($4\ Mn^{2+} + 2\ P_i$) as in the present structure are under X-ray study (Goldman, A., private communication). Several complexes of *E. coli* PPase have been crystallised by the Moscow group. In addition to a complex with a single Mn^{2+} (Harutyunyan et al., 1996), a complex containing two natural metal-activators has already been analysed (Oganessyan, V. Y., Kurilova, S. A., Vorobyova, N. N., Nazarova, T. I., Avaeva, S. M. and Harutyunyan, E. H., unpublished results). Combined efforts will certainly succeed in a better understanding of the catalytic mechanism.

This work has been supported in part by the Russian Foundation for Fundamental Studies (RFFS), grant 94-03-09-069. E. H. H. and I. P. K. thank the European Molecular Biology Laboratory for supporting visits to the Hamburg Outstation. We thank Prof. S. Avaeva for valuable discussions.

REFERENCES

Avaeva, S. M. & Nasarova, T. I. (1995) The structure and peculiarity of function of inorganic pyrophosphatase from baker's yeast, *Adv. Biol. Chem.* (*in Russian*) *26*, 42–63.

Baykov, A. A. & Avaeva, S. M. (1974) Yeast inorganic pyrophosphatase: Studies on metal binding, *Eur. J. Biochem. 47*, 57–66.

Baykov, A. A. & Shestakov, A. S. (1992) Two pathways of pyrophosphate hydrolysis and synthesis by yeast inorganic pyrophosphatase, *Eur J. Biochem. 206*, 463–470.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F. Jr, Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977) The Protein Data Bank: a computer-based archival file for macromolecular structures, *J. Mol. Biol. 112*, 535–542.

Butler, L. G. (1971) Yeast and other inorganic pyrophosphatases, in *The enzymes* (Boyer, P. D., ed.) 3rd edn, vol. 4, pp. 529–541, Academic Press, New York.

CCP4 – Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography, *Acta Crystallogr. D50*, 760–763.

Chirgadze, N. Y., Kuranova, I. P., Nevskaya, N. A., Teplyakov, A. V., Wilson, K. S., Strokopytov, B. N., Harutyunyan, E. H. & Hohne, W. E. (1991) Crystal structure of MnP_i complex of yeast inorganic pyrophosphatase at 2.35 Å resolution, *Krystallografiya (in Russian) 36*, 128–132.

Cohen, S. A., Sterner, R., Keim, P. S. & Heinrikson, R. L. (1978) Covalent structural analysis of yeast inorganic pyrophosphatase, *J. Biol. Chem. 253*, 889–897.

Cooperman, B. S., Chiu, N. Y., Bruckmann, R. H. & McKenna, G. P. (1973) Yeast inorganic pyrophosphatase. New methods of purification, assay, and crystallization, *Biochemistry 12*, 1665–1682.

Cooperman, B. S., Panackal, A., Springs, B. & Hamm, D. J. (1981) Divalent metal ion, inorganic phosphate, and inorganic phosphate analogue binding to yeast inorganic pyrophosphatase, *Biochemistry 20*, 6051–6060.

Cooperman, B. S. (1982) The mechanism of action of yeast inorganic pyrophosphatase, *Methods Enzymol. 87*, 526–548.

Cooperman, B. S., Baykov, A. A. & Lahti, R. (1992) Evolutionary conservation of active site of soluble inorganic pyrophosphatases, *Trends Biol. Sci. 17*, 262–266.

Fersht, A. (1984) *Enzyme structure and mechanism*, 2nd edn, pp. 221–247, W. H. Freeman and Co., New York.

Gonzales, M. A. & Cooperman, B. S. (1986) Glutamic acid-149 is important for enzymatic activity of yeast inorganic pyrophosphatase, *Biochemistry 25*, 7179–7185.

Hansen, G., Eifler, R. & Heitmann, P. (1972) The subunit structure of inorganic pyrophosphatase from baker's yeast, *Acta Biol. Med. Ger. 28*, 977–987.

Harutyunyan, E. H., Oganessyan, V. Y., Oganessyan, N. N., Terzyan, S. S., Popov, A. N., Rubinskiy, S. V., Vainshtein, B. K., Nazarova, T. I., Kurilova, S. A., Vorobyova, N. N. & Avaeva, S. M. (1996) Structure of inorganic pyrophosphatase from *E. coli* and its complex with a Mn^{2+} ion at 2.2 Å resolution, *Crystallographia Rep.* [English translation from *Kristallografiya (in Russian)*] *41*, 84–96.

Heikineimo, P., Salminen, T., Lahti, R., Cooperman, B. & Goldman, A. (1995) New crystal forms of *Escherichia coli* and *Saccharomyces cerevisiae* soluble inorganic pyrophosphatases, *Acta Crystallogr. D51*, 399–401.

Janson, C. A., Degani, C. & Boyer, P. D. (1979) The formation of enzyme-bound and medium pyrophosphatase and the molecular basis of the oxygen exchange reaction of yeast inorganic pyrophosphatase, *J. Biol. Chem. 254*, 3743–3749.

Jones, T. A. (1978) A graphics model building and refinement system for macromolecules, *J. Appl. Crystallogr. 11*, 268–272.

Josse, J. & Wong, S. C. K. (1971) Inorganic pyrophosphatase from *E. coli*, in *The enzymes* (Boyer, P. D., ed.) 3rd edn, vol. 4, pp. 499–527, Academic Press, New York.

Kankare, J., Neal, G. S., Salminen, T., Glumoff, T., Cooperman, B. S., Lahti, R. & Goldman, A. (1994) The structure of *E. coli* soluble inorganic pyrophosphatase at 2.7 Å resolution, *Protein Eng. 7*, 823–830.

Knight, W. B., Dunaway-Mariano, D., Ransom, S. C. & Villafranca, J. J. (1984) Investigation of the metal ion binding sites of yeast inorganic pyrophosphatase, *J. Biol. Chem. 259*, 2886–2895.

Kolakowski, L. F. Jr, Schlosser, M. & Cooperman, B. S. (1988) Cloning, molecular characterization and chromosome localization of the inorganic pyrophosphatase gene from *S. cerevisiae*, *Nucleic Acids Res. 16*, 10441–10452.

Komissarov, A. A., Makarova, I. A., Sklyankina, V. A. & Avaeva, S. M. (1985) The functionally important lysine residue in inorganic pyrophosphatase from yeast, *Bioorg. Khim. (in Russian) 11*, 1504–1509.

Konnert, J. H. & Hendrickson, W. A. (1980) A restrained-parameter thermal-factor refinement procedure, *Acta Crystallogr. A36*, 344–350.

Kuranova, I. P., Terzyan, S. S., Voronova, A. A., Smirnova, E. A., Vainstein, B. K., Höhne, W. & Hansen, G. (1983) Active site structure of the inorganic pyrophosphatase from baker's yeast based on the X-ray investigation, *Bioorg. Khim. (in Russian) 9*, 1611–1619.

Kuranova, I. P. & Sokolov, V. I. (1986) A conformational hypothesis of the trans-ligation of metals which activate pyrophosphatase and related enzymes, *Bioorg. Khim. (in Russian) 12*, 749–754.

Kuranova, I. P., Smirnova, E. A. & Chirgadze, N. Y. (1990) The growing of crystals of inorganic pyrophosphatase from yeast with metal ions and phosphate, *Krystallografiya (in Russian) 35*, 1581–1584.

Kuranova, I. P. (1992) Investigation of spatial structure of inorganic pyrophosphatase from yeast and correlation with its function, in *Structural Crystallography* (Vainshtein, B. K., ed.) vol. 1, pp. 270–289, Nauka, Moscow.

Lahti, R. (1983) Microbial inorganic pyrophosphatases, *Microbiol. Rev. 47*, 169–179.

Lahti, R., Pohjaniksa, K., Pirkaranta, T., Heikinheimo, P., Salminen, T., Meyer, P. & Heiniten, J. (1990) A site-directed mutagenesis study on *Escherichia coli* inorganic pyrophosphatase. Glutamic acid-98 and lysine-104 are important for structural integrity, whereas asparatic acids-97 and -102 are essential for catalytic activity, *Biochemistry 29*, 5761–5766.

Lahti, R., Salminen, T., Latonen, S., Heikinheimo, P., Pohjanoksa, K. & Heinonen, J. (1991) Genetic engineering of *E. coli* inorganic pyrophosphatase. Tyr55 and Tyr141 are important for the structural integrity, *Eur. J. Biochem. 198*, 293–297.

Lamzin, V. S. & Wilson, K. S. (1993) Automated refinement of protein models, *Acta Crystallogr. D49*, 129–147.

Morris, A. L., MacArthur, M. W., Hutchinson, E. G. & Thornton, J. M. (1992) Stereochemical quality of protein structure coordinates, *Proteins 12*, 345–364.

Niren, P., Hajnal, K. & Baltscheffsky, M. (1984) Purification of the membrane-bound proton-translocating inorganic pyrophosphatase from *Rhodospirillum rubrum*, *Biochim. Biophys. Acta 766*, 630–635.

Oganessyan, V. Y., Kurilova, S. A., Vorobyeva, N. N., Nazarova, T. I., Popov, A. N., Lebedev, A. A., Avaeva, S. M. & Harutyunyan, E. G. (1994) X-ray crystallographic studies of recombinant inorganic pyrophosphatase from *Escherichia coli*, *FEBS Lett. 348*, 301–304.

Otwinowski, Z. (1993) *DENZO: an oscillation data processing program for macromolecular crystallography*, Yale University, New Haven, USA.

Plaksina, E. A., Sergienko, O. V., Sklyankina, V. A. & Avaeva, S. M. (1981) Preparation of immobilized dimer and monomer of inorganic pyrophosphatase. Improvement of catalytical activity of subunits, *Bioorg. Khim. (in Russian) 10*, 357–364.

Rapoport, T. A., Hohne, W. E., Reich, J. G., Heitmann, P. & Rapoport, S. M. (1972) A kinetic model of the inorganic pyrophosphatase from bakers' yeast, *Eur. J. Biochem. 26*, 237–246.

Rapoport, T. A., Hohne, W. E., Heitman, P. & Rapoport, S. (1973) Binding of ligands to the inorganic pyrophosphatase of bakers' yeast, *Eur. J. Biochem. 33*, 341–347.

Raznikov, A. V., Sklyankina, V. A. & Avaeva, S. M. (1992) Tyrosine 89 is important for enzymatic activity of *S. cerevisiae* inorganic pyrophosphatase, *FEBS Lett. 308*, 62–64.

Read, R. J. (1986) Improved Fourier coefficients for maps using phases from partial structures with errors, *Acta Crystallogr. A42*, 140–149.

Teplyakov, A., Obmolova, G., Wilson, K. S., Ishii, K., Kaji, H., Samejima, T. & Kuranova, I. (1994) Crystal structure of inorganic pyrophosphatase from *Thermus thermophilus*, *Protein Sci. 3*, 1098–1107.

Terzyan, S. S., Voronova, A. A., Smirnova, E. A., Kuranova, I. P., Nekrasov, Y. V., Harutyunyan, E. G., Vainshtein, B. K., Hohne, W. & Hansen, G. (1984) The spatial structure of inorganic pyrophosphatase from yeast at 3 Å resolution, *Bioorg. Khimia (in Russian) 10*, 1469–1482.